

## Verfahren zur Verarbeitung von Postsachen

### Beschreibung

Die Erfindung betrifft ein Verfahren gemäß dem Oberbegriff der unabhängigen Patentansprüche. Im weitesten Sinne bezieht sich die Erfindung auf das Gebiet automatischer Briefverarbeitung und insbesondere auf Systeme, bei denen ein automatisches Adreßleseverfahren durch Videocodierung ergänzt und verbessert wird.

Systeme zum automatischen Lesen von Adressen (OCR) sind auf dem Gebiet der Briefbearbeitung wohlbekannt und z.B. in der DE 195 31 392 beschrieben. Mit modernen OCR-Briefsortieranlagen sind Bearbeitungsraten von 10 Briefen pro Sekunde, d. h. 36.000 Briefen pro Stunde und mehr erreichbar. Die Erkennungszuverlässigkeit variiert jedoch stark mit der Schriftart und Gesamtqualität der auf die Oberfläche der Briefe aufgetragenen Adreßinformation. Im Falle einer erfolgreichen Erkennung kann der betreffende Brief mit einem maschinenlesbaren Barcode versehen werden. Dieser Barcode erlaubt eine weitere mechanische Bearbeitung bis zu einer beliebigen gewünschten Sortierordnung. Insbesondere ermöglicht die Verwendung von Barcodes eine Sortierung der Briefe bis zur Sortierebene des Postgangs, bei der Briefe entsprechend der Reihenfolge ihrer Verteilung durch den Zusteller sortiert werden.

Da die Erkennungsraten der automatischen Lesesysteme sehr variieren, ist es notwendig, diese durch verschiedene Formen von manueller Intervention zu unterstützen. Die einfachste Intervention ist Zurückweisung nicht automatisch lesbarer Briefe, und die Durchführung eines Handsortierverfahrens. Allerdings sind die dadurch entstehenden Kosten mit zunehmenden Arbeitskosten unökonomisch hoch. Hinzu kommt, daß solche handsortierte Post auch zu späteren Zeitpunkten nicht mehr ohne weiteres mechanisch weitersortiert werden kann, so daß zwei voneinander getrennte Ströme von Sendungen erzeugt werden, die wiederum manuell zu einem bestimmten Zeitpunkt verschmolzen werden müssen.

Um diese Nachteile manueller Sortierung von OCR-zurückgewiesenen Sendungen zu vermeiden, sind verschiedene Verfahren für eine manuelle Codierung von Poststücken entwickelt worden. Alle diese Verfahren verwenden Interventionen

Um die Codierproduktivität zu erhöhen und/oder die Angabe von sämtlichen Adreßelementen, d. h. ZIP/PC, Straße/Postfach, Adressat/Postfach, Adressat/Firma, zu ermöglichen, sind im Stand der Technik verschiedene Methoden entwickelt worden. Im wesentlichen sind dies:

#### Vorschau-Codierung

Bei der Vorschau -Codierung erfolgt ein simultaner Display der Bilder von zwei Sendungen erfolgt; eines über dem anderen. Hierbei ist das niedrigere Bild das aktive, d. h. dessen Daten codiert werden. Nach einem geeigneten Training ist es Operatoren möglich die Information auf dem unteren Bild zu codieren, während sie bereits die Adreßinformation vom oberen Bild aufnehmen. Das obere Bild wird anschließend aktiv und der Prozeß wird fortgesetzt. Mit der Vorschau-Codierung ist es möglich, durch eine vollständige Überlappung der kognitiven und motorischen Funktionen beim Codieren aufeinanderfolgender Abbilder die Operator-Produktivität zu verdoppeln.

#### Extraktionscodierung

Da bei dem praktisch erreichbaren On-line-Verzögerungs-Zeiten lediglich die ZIP/PC Adresselemente verläßlich vom Operator eingegeben werden können, werden beim Extraktionscodieren bestimmte Schlüsselbestandteile der sich auf die Straße beziehenden Adressbestandteils eingegeben. Üblicherweise basiert die Extraktionscodierung auf speziell entwickelten Regeln, bei denen ein Codefester Länge, als Zugangsschlüssel zu einem Adressverzeichnis verwendet wird. Zum Beispiel verwendet die Royal Mail eine Extraktionsformel, die auf den ersten drei und den letzten zwei Buchstaben basiert. Dabei müssen Sonderregeln vom Operator auswendig gelernt werden, um überflüssige Adressinformationen zu vermeiden und bestimmte unterscheidende Merkmale wie z.B. Richtungen zu berücksichtigen, z.B. Ost, West oder Kategorien z.B. Street, Lane, Road.

Extraktionscodierung besitzt trotz ihrer gewissen Effektivität einige größere Nachteile; insbesondere komplexe Extraktionsregeln, die häufig die Berücksichtigung des Endes eines Straßennamens erfordern, während diese Bestandteile üblicherweise am unklarsten geschrieben sind. Außerdem findet man eine signifikant hohe Rate von nicht eindeutigen Extraktionen bei denen mehrere Einträge in einem Verzeichnis dem Extraktionscode entsprechen, so daß keine eindeutige Sortierentscheidung getroffen werden kann. Desweiteren ist zu berücksichtigen, daß die Eingabeproduktivität der Operatoren vermindert wird, sobald vom Operator Entscheidungen getroffen werden müssen, statt einer einfachen repetitiven Tastatureingabe.

Weiterverarbeitung mit nichtgelesenen Adressen versehenen Sendungen und eine entsprechend komplexe Logistik verlangt.

Aufgabe der vorliegenden Erfindung ist es, eine möglichst hohe on-line-Codierung von Briefsendungen und eine möglichst geringe Rate von Sendungen mit nicht erfolgreich gelesenen Adressen zu erreichen. Erfindungsgemäß wird diese Aufgabe durch die Merkmale der unabhängigen Patentansprüche gelöst.

Die Erfindung ermöglicht eine verbesserte Integration von automatischen Lesesystemen und Video-Codierung. Weiter ermöglicht die Erfindung die effektive Verwendung einer Extraktionscodierung in integrierten automatischen und Video-Codiersystem, insbesondere auch eine Vereinfachung des Entscheidungsproblems für die Operatoren bei der Adresscodierung. Ein weiterer Vorteil des erfindungsgemäßen Verfahrens ist, daß zusätzliche Sortierinformationen, z.B. solche, die sich auf die Namenszeile der Adresse beziehen, effektiv ausgewertet werden können. Ein weiterer Vorteil der Erfindung besteht darin, daß Inkonsistenzen, die sich z.B. bei der Extraktionscodierung oder durch Fehler bei der Beschriftung von seiten des Absenders ergeben, effektiv behandelt werden können. Das Verfahren ermöglicht weiter eine einfache Integration von on-line und off-line-Codierung sowie des Vorschau-Codierungsverfahrens.

Vorteilhafte Ausführungsformen der Erfindung sind den abhängigen Ansprüchen und der Beschreibung zu entnehmen.

Im folgenden wird die Erfindung anhand von Zeichnungen beispielhaft dargestellt. Im einzelnen zeigen:

Fig. 1 eine schematische Darstellung einer Vorrichtung zur Ausführung des Verfahrens

Fig. 2 und 3 eine Übersicht über den Datenfluß gemäß der Erfindung

Fig. 1 zeigt eine schematische Darstellung einer Briefverteilanlage mit der das erfindungsgemäße Verfahren ausgeführt werden kann. Der OCR-Briefsortierer 100 besteht aus einer Zuführeinrichtung 110, die sukzessive Sendungen aus einem Magazin 111 abzieht und mit ca. 10 Sendungen pro Sekunde zu einem hochauflösenden Videoscanner 120 transportiert. Anschließend werden die Sendungen in einer Verzögerungsstrecke 121 transportiert. Die Sendungen weisen üblicherweise auf ihrer Oberfläche Adressinformationen auf. Im OCR-Prozessor 130 erfolgt eine Auswertung der Adressinformationen der Abbilder der Sendungen die vom Videoscanner 120 gewonnen wurden. Im Fall einer vollständigen

Extraktionscodes für die Straße „Hell“ sowie die Hausnummer „8“. Statt eines zweigeteilten Displays sind auch andere Formen möglich. Diese Eingabeinformationen werden benutzt, um im Adressenverzeichnis 134 passende Einträge zu identifizieren. Für den Fall, daß eine eindeutige Zuordnung zwischen der eingegebenen Information und einem Eintrag im Adressenverzeichnis 134 gefunden wurde, ist damit grundsätzlich eine vollständige Auswertung der Adressinformation des betreffenden Abbildes erfolgt. Allerdings ist für einen gewissen Prozentsatz von Eingaben eine solche Eindeutigkeit nicht erreichbar, da eine Anzahl verschiedener Adresseinträge der codierten Adressinformation entsprechen.

Gemäß Fig. 3 erfolgt eine Entscheidung 300, ob eine vollständige Auswertung der Adressinformation eines Abbildes bei der Video-Codierung erfolgt ist. Ist die Entscheidung positiv (Yes) kann die entsprechende Sendung entweder mit einem Barcode versehen werden, wenn die Verzögerungszeit ausgereicht hat, um eine On-line-Video-Codierung durchzuführen oder es wird eine entsprechende Verknüpfung der TID und eine darauf basierende Barcodierung vorgenommen. In jedem Fall kann die entsprechende Sendung mit üblichen Mitteln weiter sortiert werden. Bei einer negativen Entscheidung (No) erfolgt erfindungsgemäß eine weitere automatische Auswertung unter Verwendung der Ergebnisse der Video-Codierung, d.h. bei dieser weiteren automatischen Auswertung stehen dem OCR-Prozessor neben den auf der Abbild dargestellten Informationen die durch die Video-Codierung gewonnenen Informationen zur Verfügung; im obigen Beispiel das Tripel „4432“, „Hell“, „8“. Dies ist in Fig. 3 symbolisch durch den Inhalt des Kreises 310 zum Ausdruck gebracht. Anschließend erfolgt eine Entscheidung 320, ob nunmehr eine vollständige Auswertung des betreffenden Bildes stattgefunden hat. Im positiven Fall (Yes) wird die entsprechende Information zum weiteren Sortieren der Sendung verwendet analog wie nach einer positiven Entscheidung beim Entscheidungspunkt 300. Ist die Entscheidung negativ (No) erfolgt eine weitere Video-Codierung unter Verwendung der Ergebnisse der weiteren automatischen Auswertung. Vorzugsweise werden dabei dem Operator eine Anzahl von Alternativen zur Selektion vorgeführt, aus denen eine Selektion vorzunehmen ist.

Vorzugsweise werden die beiden letzten Schritte des Verfahrens nämlich die weitere automatische Auswertung sowie die weitere Video-Codierung nicht mehr on-line sondern off-line erfolgen, weil die zur Verfügung stehende Verzögerungszeit zu gering ist für eine on-line-Ausführung.

Eintragungen im Adressverzeichnis vor. Bei geeigneter Extraktionscodierung sind dabei in über 90% der Fälle lediglich zwei Adresseinträge vorhanden.

3. Weitere Video-Codierung derjenigen Abbilder, für die bei den vorhergehenden Phasen keine vollständige Auswertung möglich war. Vorzugsweise werden Abbilder der nicht ausgewerteten Adressen dem Operator präsentiert, zusammen mit den Ergebnissen der vorhergehenden Video-Codierung und der automatischen Auswertung der zweiten Phase zwei. Vorzugsweise hat der Operator hierbei lediglich eine Selektion aus vorgegebenen Alternativen vorzunehmen.

Bei der weiteren automatischen Auswertung liegt nunmehr zusätzliche Kontextinformationen vor und zwar ist die Anzahl der zu berücksichtigen Adresseinträge beschränkt durch die Anzahl der aus der Extraktionscodierung erhaltenen Kandidaten. Es kann davon ausgegangen werden, daß unter diesen Kandidaten die richtige Adresse zu finden ist. Zusätzlich ist üblicherweise die Hausnummer bekannt.

Bei der weiteren Video-Codierung solcher Adressinformationen, die in den vorhergehenden Phasen nicht vollständig ausgewertet wurden, werden vorzugsweise nicht eindeutige Ergebnisse der Extraktionscodierung oder zusätzliche sortierrelevante Informationen, die sich auf der Adressatenlinie befinden, bearbeitet. Vorzugsweise werden dem Operator sukzessive Bilder der Sendungsoberfläche präsentiert, wobei in einem Fenster die Auswertungsalternativen der Adresse oder des Adressaten dargestellt sind. Die Auswahl kann entweder durch Tastatureingabe einer Auswahlnummer, per Maus oder über einen Sprachprozessor erfolgen.

Bei einer bevorzugten Ausführungsform der Erfindung erfolgt eine Auswertung eines ersten Bestandteils der Adressinformation und eine Auswertung eines zweiten Bestandteils der Adressinformation und eine Prüfung der Ergebnisse dieser Auswertungen in Hinblick auf gegenseitige Konsistenz. Der erste Bestandteil der Adressinformation kann insbesondere der ZIP- oder Postcode sein, der zweite eine Straßenangabe oder eine Hausnummer. Inkonsistenzen zwischen diesen beiden Bestandteilen können z.B. durch Lesefehler oder durch eine Falschangabe des ZIP- oder Postcodes vorkommen. Bei einer Falschangabe werden ausgehend von den ersten drei oder vier Ziffern des ZIP oder Postcodes eine Anzahl von Alternativen für den Straßennamen gewonnen. Vom Operator erfolgt eine Extraktionscodierung des zweiten Bestandteils der Adressinformation, woraus

Patentansprüche

1. Verfahren zur Bearbeitung von Sendungen in einem automatischen Adreslesesystem, wobei ein Abbild der Adressinformation aufweisenden Oberfläche jeder Sendung gewonnen und einer ersten automatischen Auswertung zugeführt wird, und nicht vollständig ausgewertete Adressinformation einer ersten Video-Codierung zur Auswertung zugeführt wird,  
**dadurch gekennzeichnet,**  
daß die Adressinformation derjenigen Abbilder, die bei der Video-Codierung nicht vollständig ausgewertet worden sind, unter Verwendung der Ergebnisse der Video-Codierung einer weiteren automatischen Auswertung zugeführt werden.
2. Verfahren zur Bearbeitung von Sendungen nach Anspruch 1,  
dadurch gekennzeichnet,  
daß die Adressinformation diejenigen Abbilder, die bei der weiteren automatischen Auswertung nicht vollständig ausgewertet worden sind, einer weiteren Video-Codierung zur Auswertung unter Verwendung der Ergebnisse der weiteren automatischen Auswertung zugeführt werden.
3. Verfahren nach Anspruch 1 oder 2,  
dadurch gekennzeichnet,  
daß bei der ersten Video-Codierung eine Extraktionscodierung gemäß vorgegebener Extraktionsregeln erfolgt.
4. Verfahren nach Anspruch 2 oder 3,  
dadurch gekennzeichnet,  
daß bei der weiteren Video-Codierung eine Selektionscodierung derart erfolgt, daß aus einer Anzahl von alternativen Auswertungsergebnissen eine Selektion vorgenommen wird.
5. Verfahren nach Anspruch 4,  
dadurch gekennzeichnet,  
daß die alternativen Auswertungsergebnisse aus den nicht eindeutigen Ergebnissen der Extraktionscodierungen gebildet werden.

13. Vorrichtung zur Durchführung des Verfahrens nach Ansprüchen 1 bis 12, mit einem automatisches Adresslesesystem, daß
- eine Vorrichtung zur Gewinnung von Abbildern von Sendungen,
  - eine Vorrichtung zur automatischen Auswertung von Adressinformation
  - eine Vorrichtung zur Video-Codierung von Adressinformation
  - einen Imagecontroller zur Steuerung des Datenflusses zwischen der Vorrichtung zur automatischen Auswertung und der Vorrichtung zur Video-Codierung aufweist, wobei die Vorrichtung zur Video-Codierung eine Anzahl von Video-Codieplätzen aufweist,
- dadurch gekennzeichnet,
- daß der Imagecontroller derart ausgebildet ist, daß von der Video-Codierung nicht vollständig ausgewertete Adressinformation unter Verwendung der Ergebnisse der Video-Codierung der Vorrichtung zu einer weiteren automatischen Adressauswertung zugeführt werden.
14. Vorrichtung nach Anspruch 13,
- dadurch gekennzeichnet,
- daß eine Vorrichtung vorgesehen ist, um Sendungen, deren Adressinformationen nicht vollständig on-line ausgewertet worden ist, mit einer Identifizierungsinformation (TID) zu versehen.

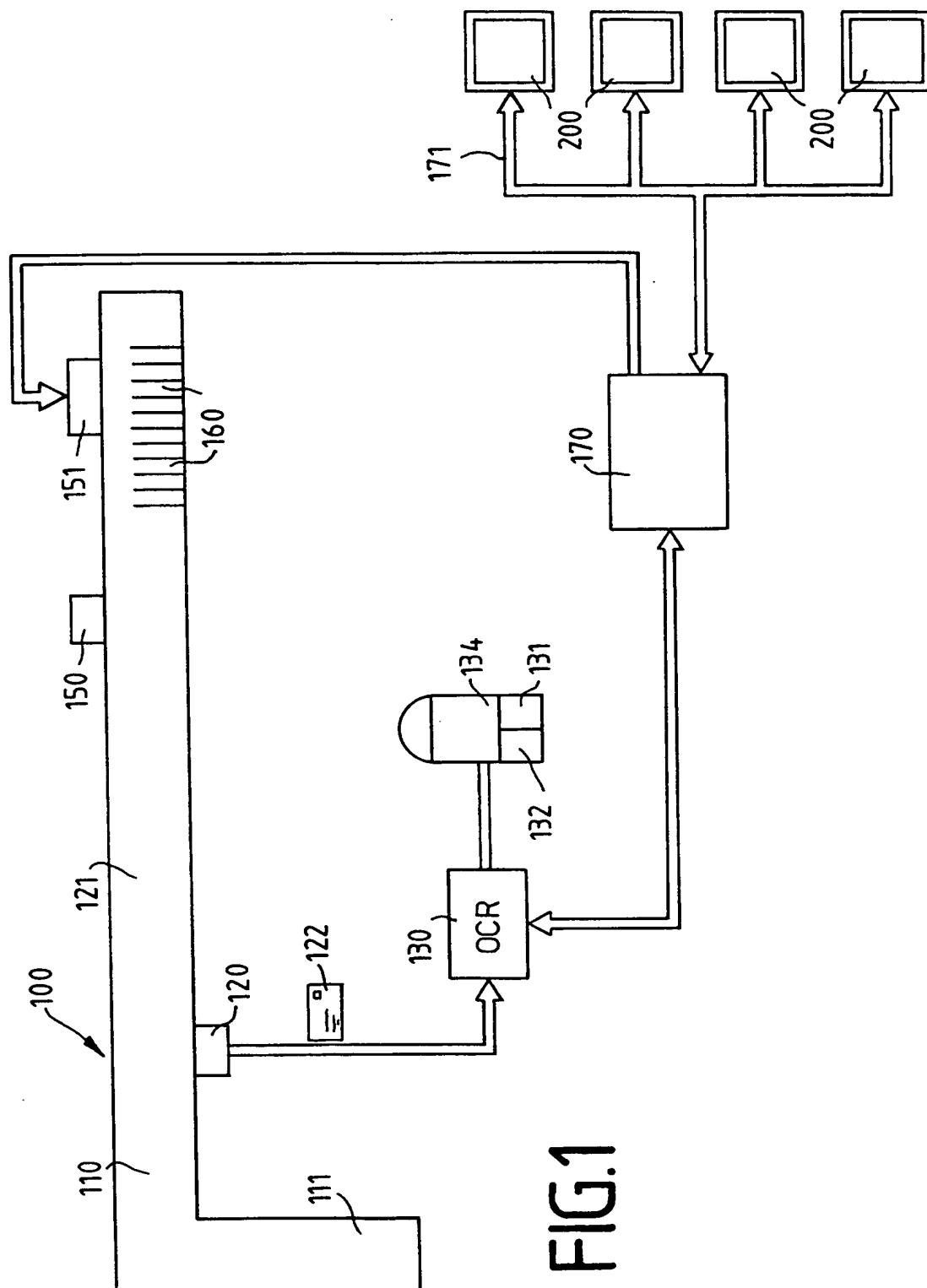




FIG. 2

